Plagiarism Checker using Machine Learning

Prof. Vanita Buradkar[1], Rutuja Chawale[2], Janhvi Rangankar[3],
Saptashrungi Chaudhari[4], Vijaya Zade[5]

Assistant Professor, Department of Computer Science & Engineering[1]
Students, Department of Computer Science & Engineering[2,3,4,5]
Rajiv Gandhi College of Engineering, Research & Technology, Chandrapur, Maharashtra, India.

Abstract: *"Plagiarism, the act of taking the writings of another person and passing them off as one's own. The fraudulence is closely related to forgery and piracy-practices generally in violation of copyright laws." Encyclopedia Britannica . Plagiarism can be considered as one of the electronic crimes, like (computer hacking, computer viruses, spamming, phishing, copyrights violation and others crimes). Plagiarism defined as the act of taking or attempting to take or to use(whole or parts) of another person's works, without referencing or citation him as the owner of this work. It may include direct copy and paste, modification or changing some words of the original information from the internet books, magazine, newspaper, research, journal, personal information or ideas. According to the MerriamWebster On Line Dictionary, to "plagiarize" means 1)to steal and pass off (the ideas or words of another) as one's own 2)to use (another's production) without crediting the source 3)to commit literary theft 4)to present as new and original an idea or product derived from an existing source. "Technology has been both a miracle and a curse in terms of plagiarism. No doubt, it has become easier to find the required information and copy it. Since people often do that without attribution, it has also become easier to identify and deal with plagiarism." With free plagiarism checker tools that can search billions of documents, and find matches even if they are only a few words in length, finding plagiarism has become as easy as detecting information in Google. It is now only a matter of merely processing your query and giving you the results. "Plagiarism definition is actually straightforward. When you use someone else's work without crediting them, it is seen as stealing their intellectual property. Just like theft, the penalties for plagiarized work are also severe all over the world. The real problem is that most people are not even aware of what they are doing." By looking the above parameter we are proposed a mechanism which will helps to workout for plagiarism checker application which will helps the peoples to find the identical content with certain papers so that the proposed mechanism will helps the people to go for there unique content*

Keywords: Plagiarism

REFERENCES

[1] Sanjay Goel, Deepak Rao et al.: Plagiarism and its Detection in Programming Languages, December 15, 2005.

[2] http://www.php.net , (accessed: 23/10/2009)

[3] http://docs.moodle.org/en/Talk:About_Moodle, (accessed: 23/10/2009)

[4] Manuel Freire, Manuel Cebrian and Emilio del Rosal: An Integrated Source Code Plagiarism Detection Environment, Escuela Politecnica Superior, Universidad Autonoma de Madrid, 28049 Madrid, Spain.

[5] Shared Information and Program Plagiarism Detection, Xin Chen, Brent Francia, Ming Li, Brian Mckinnon, Amit Seker, University of California, Santa Barbara, December 13, 2003.

[6] Paul Clough: Plagiarism in natural and programming languages: an overview of current tools and technologies, Department of Computer Science, University of Sheffield, July 2000.

[7] JPlag: Finding plagiarisms among a set of programs, University at Karlsruhe D76128 Karlsruhe, Germany

[8] Manuel Freire, Manuel Cebrifn and Emilio del Rosal Escuela: AC: An Integrated Source Code Plagiarism Detection Environment, Polit_ecnica Superior, Universidad Aut_onoma de Madrid.

[9] http://software.bioinformatics.uwaterloo.ca/SID/servlets/Index Page , (accessed: 18/11/2009)